Verständnis für die neue generative KI-Inhaltskontrolle und die erweiterte Abdeckung durch DLP-KI-Tools

Inhalt

Einleitung

Dieses Dokument beschreibt die neue generative KI-Inhaltskontrolle und die Erweiterung der Abdeckung der DLP-KI-Tools für Umbrella.

Überblick

Wir freuen uns, die allgemeine Verfügbarkeit von Generative AI Content Control bekannt geben zu können.
Mit dieser Funktion können Sie von ChatGPT generierte Inhalte überwachen und bei Bedarf blockieren.

Wir freuen uns auch , Ihnen mitteilen zu können, dass wir den Umfang unserer Echtzeit-DLP-Abdeckung für Generative KI-Tools erweitert haben. Anfänglich auf ChatGPT beschränkt, unterstützen wir nun alle 70 KI-Tools in unserer kürzlich veröffentlichten Generative KI-Anwendungskategorie. Diese deutliche Erweiterung ermöglicht es Ihnen, den Anwendungsfall der sicheren Nutzung von KI zu erweitern und bietet eine umfassendere und robustere Lösung für den generativen Schutz der KI-Nutzung.

Wie kann DLP dazu beitragen, den von ChatGPT generierten Inhalt zu kontrollieren?

DLP kann Organisationen bei der Kontrolle generierter Inhalte unterstützen, indem ChatGPT-Antworten mithilfe von Echtzeit-DLP-Richtlinien gescannt werden. Mit dieser Version können Sie

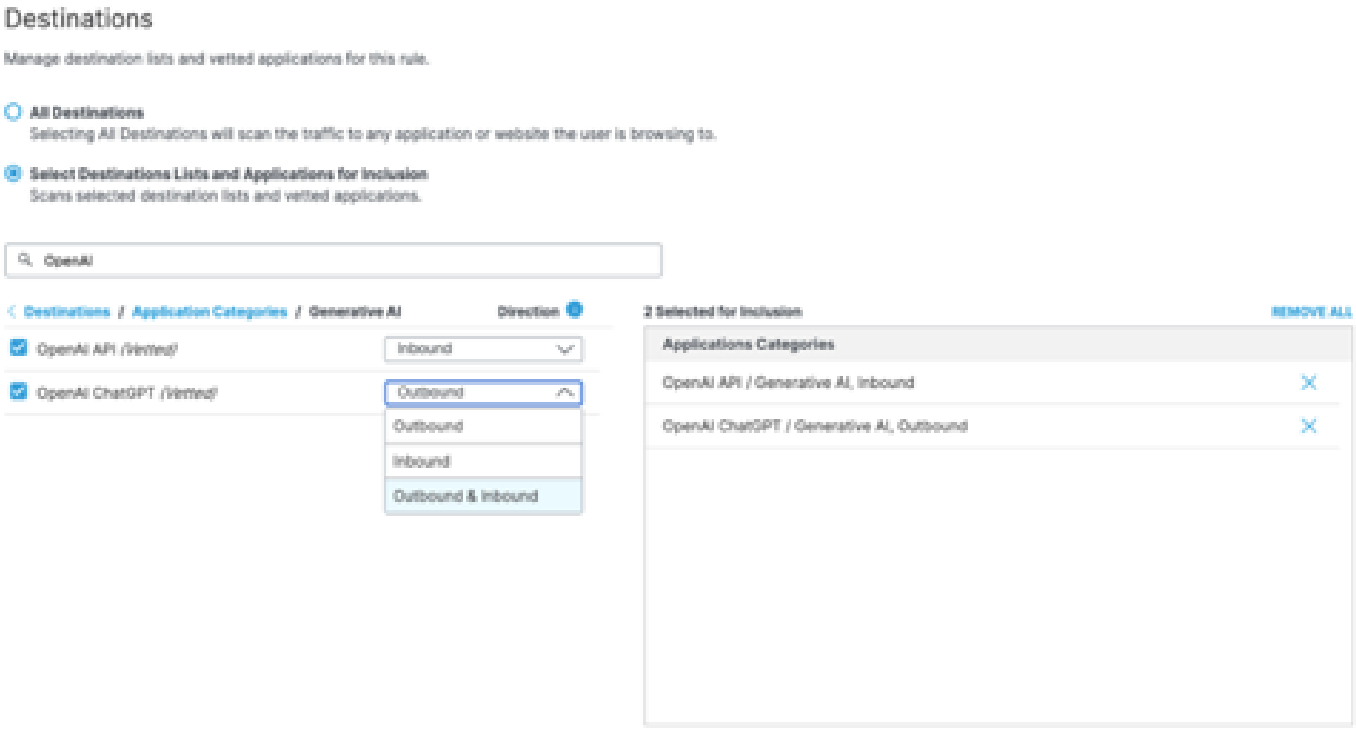
festlegen, dass ChatGPT-Antworten (d. h. eingehender Datenverkehr) nach allen Arten von generierten Inhalten durchsucht werden, die überwacht oder blockiert werden sollen.

Warum sollten KI-generierte Inhalte kontrolliert werden?

Die Verwendung von durch künstliche Intelligenz generierten Inhalten birgt für Unternehmen aus verschiedenen Gründen Risiken, darunter Urheberrechtsverletzungen, ungenaue Informationen, fehlerhaften Code usw.
So können Sie beispielsweise verhindern, dass Benutzer den von AI generierten Quellcode verwenden, um die Verwendung von urheberrechtlich geschütztem oder unsicherem Code zu verhindern, während andere Benutzer die Verwendung von durch AI generierten Gerichtsurteilen aus Angst vor der Eingabe falscher Informationen verhindern möchten.

Wie kann ich DLP-Scanning auf ChatGPT-Antworten anwenden?

Im Allgemeinen scannt Real-Time DLP ausgehenden Webdatenverkehr, wie ChatGPT-Aufforderungen, um Datenlecks zu verhindern. Mit dieser Version können wir auch eingehenden Datenverkehr scannen, indem wir die Richtung des Datenverkehrs auswählen, der von der Echtzeit-DLP gescannt wird, d. h. eingehender Datenverkehr, ausgehender Datenverkehr oder beides. Diese Fähigkeit ist derzeit nur für ChatGPT (sowohl chatbot und API). Bei Auswahl von Eingehender Datenverkehr wird ChatGPT-Antworten gescannt.



23281122679316

Was ist die Anwendungskategorie "Generative KI" in DLP?

Vor dieser Version enthielten die Zielkriterien in den Echtzeit-DLP-Regeln eine endliche Liste mit 20 Anwendungen, die ausgewählt werden konnten. Mit dieser Version ermöglicht Real-Time DLP

Kunden die Auswahl einer unserer 38 Anwendungskategorien, einschließlich Generative AI, oder einer der ± 4.600 verfügbaren kontrollierbaren Anwendungen, die in ihnen kategorisiert sind. Die Anwendungskategorie "Generative KI", die erst vor wenigen Monaten mit 20 Apps auf den Markt kam, umfasst jetzt 70 Apps, und wir sind entschlossen, diese Kategorie kontinuierlich mit erstklassigen KI-Tools zu aktualisieren.

Kann eine SvD-Regel auf die gesamte Anwendungskategorie der generativen KI angewendet werden?

Ja, eine Echtzeit-SvD-Regel kann auf eine ganze Kategorie oder eine Teilmenge der darin enthaltenen Anwendungen angewendet werden.

Wo finde ich zugehörige Dokumente?

- So kontrollieren Sie die Scanrichtung, um ChatGPT-Antworten zu überwachen oder zu blockieren:
  Hinzufügen einer Echtzeitregel zur Richtlinie zum Schutz vor Datenverlust
- Hier erfahren Sie, wie Sie überprüfen, ob eine chatGPT-Eingabeaufforderung oder eine chatGPT-Antwort blockiert wurde, und wie Sie die Scanrichtung überprüfen. Bericht zum Schutz vor Datenverlust
- Um alle Anwendungskategorien zu überprüfen, die jetzt in den Echtzeit-SvD-Policy-Regeln verfügbar sind, aktivieren Sie hier: Anwendungskategorien

Haben wir vor, bei der kommenden Cisco Live in Amsterdam eine Ankündigung bezüglich dieser hochinteressanten Anwendungsfälle für den Schutz vor generativer KI zu machen?

Ja, wir werden am Dienstag, 6. Februar, 15:00 bis 16:30 Uhr MEZ, eine Breakout-Session mit dem Titel Protecting Your Sensitive Data from Generative AI Usage in Cisco Live Amsterdam abhalten.

Bitte nehmen Sie Platz!

Informationen zu dieser Übersetzung

Cisco hat dieses Dokument maschinell übersetzen und von einem menschlichen Übersetzer editieren und korrigieren lassen, um unseren Benutzern auf der ganzen Welt Support-Inhalte in ihrer eigenen Sprache zu bieten. Bitte beachten Sie, dass selbst die beste maschinelle Übersetzung nicht so genau ist wie eine von einem professionellen Übersetzer angefertigte. Cisco Systems, Inc. übernimmt keine Haftung für die Richtigkeit dieser Übersetzungen und empfiehlt, immer das englische Originaldokument (siehe bereitgestellter Link) heranzuziehen.